Yu. V. Averboukh, Markov approximations of nonzero-sum differential games, *Vestn. Udmurtsk. Univ. Mat. Mekh. Komp. Nauki*, 2020, Volume 30, Issue 1, 3–17

DOI: https://doi.org/10.35634/vm200101

VESTNIK UDMURTSKOGO UNIVERSITETA. MATEMATIKA. MEKHANIKA. KOMP'UTERNYE NAUKI
MATHEMATICS 2020. Vol. 30. Issue 1. Pp. 3–17.

MSC2010: 91A23, 91A10, 91A05

© *Yu. V. Averboukh*

MARKOV APPROXIMATIONS OF NONZERO-SUM DIFFERENTIAL GAMES

The paper is concerned with approximate solutions of nonzero-sum differential games. An approximate Nash equilibrium can be designed by a given solution of an auxiliary continuous-time dynamic game. We consider the case when dynamics is determined by a Markov chain. For this game the value function is determined by an ordinary differential inclusion. Thus, we obtain a construction of approximate equilibria with the players' outcome close to the solution of the differential inclusion. Additionally, we propose a way of designing a continuous-time Markov game approximating the original dynamics.

Keywords: nonzero-sum differential games, approximate Nash equilibria, Markov games, differential inclusion.

DOI: 10.35634/vm200101

Introduction

Nonzero-sum differential games arise in different areas of science including economics, finance, engineering and ecology. There are several solution concepts examined for nonzero-sum games. The most popular concept is the Nash equilibrium. It refers to the case when the players play noncooperatively and choose the strategies simultaneously. The most tempting approach to the study of Nash equilibrium for nonzero-sum differential games is the dynamic programming [3, 9]. It reduces the original game-theoretical problem to the system of Bellman equations which are for this case first-order PDEs. If a solution of the system of Bellman equations exists, it should provide a so called subgame-perfect Nash equilibrium. However, up to now there are no general existence results for this system of PDEs. There are only few results concerning games with simple dynamics [4–6].

The alternative way is given by punishment strategies which extend the concept of folk theorems to the differential games. In this case the players choose a trajectory, move along it, whereas any individual deviations are punished by other players [7, 12, 14, 20]. Using the punishment strategies technique one can prove the existence of a Nash equilibrium. On the other hand, this approach leads to multiple equilibria. Nowadays, there is no natural way to select a proper solution from this set. Additionally, the threats required to realize punishment strategies often look incredible.

Notice that the games with continuous-time stochastic dynamics are simpler. The existence results for the system of Bellman equations are proved for wide classes of stochastic differential games [11, 17, 18] and for Markov games [16]. Notice that for a stochastic differential game the system of Bellman equations is the system of second-order parabolic PDEs, whereas for a Markov game this system is reduced to the system of ordinary differential inclusions. Thus, the natural idea is to construct an approximate equilibrium for a nonzero-sum differential game based on the solution of the game with the dynamics given by a stochastic process. The general theory which assumes that the solution of a nonzero-sum continuous-time stochastic game is known was developed in [2]. The proposed construction assumes that the players can observe a common public signal that is produced by a model stochastic game. The important particular case is the construction of an approximate Nash equilibrium based on the solution of a stochastic differential game also examined in [2].

The aim of the paper is to construct an approximate Nash equilibrium based on the solution of a Markov game, i. e., continuous-time stochastic dynamical game with the dynamics determined by a Markov chain. Notice that a value function of a Markov game can be obtained as a solution of the differential inclusion which plays the role of the system of Bellman PDEs. We prove that given a solution of this inclusion, one can construct an approximate Nash equilibrium for the original game in the class of stochastic strategies with memory. Additionally, we provide the method of approximation of the original nonzero-sum game by the nonzero-sum Markov game.

The paper is organized as follows. First, we recall the general construction of approximate Nash eqilibria based on the solution of continuous-time stochastic games. Then in Section 2 we show that given a solution of the differential inclusion playing the role of the Bellman equation for a nonzero-sum Markov game, one can construct an approximate Nash equilibrium in the original game. In this case an error rate is estamated by the distance between the original and Markov games. Finally, in Section 3 we introduce the construction of a Markov game approximating the original differential game.

§ 1. Stochastic strategies for nonzero-sum differential games

We examine a nonzero-sum differential game with the dynamics

$$\frac{d}{dt}x(t) = f_1(t, x(t), u(t)) + f_2(t, x(t), v(t)), \ \ t \in [t_0, T], \ \ x(t) \in \mathbb{R}^d, \ \ u(t) \in U, \ \ v(t) \in V. \quad (1.1)$$

Here $u(t)$ (respectively, $v(t)$) stands for the control of the first (respectively, second) player; U and V are sets playing the role of control spaces for the players. It is assumed that player i tries to maximize $\gamma_i(x(T))$.

Let us informally discuss the strategies used in the paper. The strong formalization is given below in Definition 1. We consider the case when the players form their controls in a stochastic way using a public signal that a stochastic process observed by both players at each time. Additionally, we assume that the players' controls depend on the history, i. e., they depend on the trajectory $x(\cdot) \in C([t_0, T]; \mathbb{R}^d)$ in the nonanticipative way. This leads to the concept of public-signal stochastic strategies with memory proposed in [2].

To introduce this we need some additional notations. If $s, r \in [0, T]$, then denote by $\mathbb{F}_{s,r} \triangleq \mathcal{B}(C([s, r], \mathbb{R}^d))$, where $\mathcal{B}(\mathcal{X})$ stands for the Borel σ-algebra on a metric space $(\mathcal{X}, \varrho)$. Further, recall that if $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]})$ is a filtered measurable space [19], then the process $Z(\cdot)$ taking values in some metric space $(\mathcal{X}, \varrho)$ is said to be $\{\mathcal{F}_t\}_{t\in[t_0,T]}$-adapted if, for any $t \in [0, T]$, the mapping

$$\Omega \ni \omega \mapsto Z(t, \omega) \in \mathcal{X}$$

is $\mathcal{F}_t$-measurable [19, D31]. Additionally, the process $Z(\cdot)$ is called $\{\mathcal{F}_t\}_{t\in[t_0,T]}$-progressively measurable if, for any $t \in [0, T]$, the mapping

$$[0, t] \times \Omega \ni (s, \omega) \mapsto Z(s, \omega) \in \mathcal{X}$$

is measurable with respect to $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ [19, D45]. Hereinafter, $\otimes$ stands for the product of σ-algebras. Obviously, the progressively measurable process is adapted. If the $\{\mathcal{F}_t\}_{t\in[t_0,T]}$-adapted process $Z(t)$ takes values in the Euclidean space, then one can construct its $\{\mathcal{F}_t\}_{t\in[t_0,T]}$-progressively measurable modification [19, T46].

Definition 1. We say that a 6-tuple $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, u_{x(\cdot)}, v_{x(\cdot)}, P_{x(\cdot)})$ is a profile of public signal stochastic strategies if

(i) $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]})$ is a measurable space with filtration;

(ii) for every $x(\cdot) \in C([t_0, T]; \mathbb{R}^d)$, $P_{x(\cdot)}$ is a probability on $\mathcal{F}$;

(iii) for every $x(\cdot) \in C([t_0, T]; \mathbb{R}^d)$, $u_{x(\cdot)}$ (respectively, $v_{x(\cdot)}$) is a $\{\mathcal{F}_t\}_{t\in[t_0,T]}$-progressively measurable process taking values in U (respectively, V);

(iv) if $x(t) = y(t)$ for each $t \in [t_0, r]$, then

- for every $A \in \mathcal{F}_r$, $P_{x(\cdot)}(A) = P_{y(\cdot)}(A)$,
- for every $t \in [t_0, r]$, $u_{x(\cdot)}(t) = u_{y(\cdot)}(t)$, $v_{x(\cdot)}(t) = v_{y(\cdot)}(t)$ $P_{x(\cdot)}$-a.s.

(v) for any r, the restrictions of functions $(x(\cdot), t, \omega) \mapsto u_{x(\cdot)}(t, \omega)$, $(x(\cdot), t, \omega) \mapsto v_{x(\cdot)}(t, \omega)$ on $C([t_0, T]; \mathbb{R}^d) \times [t_0, r] \times \Omega$ are measurable with respect to $\mathbb{F}_{t_0,T} \otimes \mathcal{B}([t_0, r]) \otimes \mathcal{F}_r$;

(vi) for any $A \in \mathcal{F}$, the function $x(\cdot) \mapsto P_{x(\cdot)}(A)$ is measurable with respect to $\mathbb{F}_{t_0,T}$.

Let us briefly explain this definition. We assume that the players' controls depend both on a random signal and the trajectory. Additionally, the probabilities of the signal are determined by the trajectory. Conditions (i)–(iii) and (vi) provide measurability properties of the introduced objects. Conditions (iv) and (v) state that the dependence of the controls and the probability of the signal on the trajectory is nonanticipating.

Now let us define the motion produced by the profile of public signal stochastic strategies. Notice that since we consider stochastic strategies a realization should be a stochastic process and include also a probability that is consistent with the family of probabilities $P_{x(\cdot)}$.

Definition 2. Let $t_0 \in [0, T]$, $x_0 \in \mathbb{R}^d$, $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, P_{x(\cdot)}, u_{x(\cdot)}, v_{x(\cdot)})$ be a profile of public-signal correlated strategies on $[t_0, T]$. We say that a pair $(X(\cdot), P)$ is a realization of the motion generated by $\mathfrak{w}$ and initial position (t_0, x_0) if

(i) P is a probability on $\mathcal{F}$;

(ii) $X(\cdot)$ is a $\{\mathcal{F}_t\}_{t\in[t_0,T]}$-adapted process taking values in $\mathbb{R}^d$;

(iii) $X(t_0) = x_0$ P-a.s.;

(iv) for P-a.e. $\omega \in \Omega$,

$$\frac{d}{dt}X(t, \omega) = f_1(t, X(t, \omega), u_{X(\cdot,\omega)}(t, \omega)) + f_2(t, X(t, \omega), v_{X(\cdot,\omega)}(t, \omega)).$$

(v) $P_{x(\cdot)} = P(\cdot|X(\cdot) = x(\cdot))$, i. e., given $A \in \mathcal{F}$,

$$P(A) = \int_{C([t_0,T];\mathbb{R}^d)} P_{x(\cdot)}(A)\chi(d(x(\cdot))),$$

where χ is a probability on $\mathcal{B}(C([t_0, T]; \mathbb{R}^d))$ defined by the rule: for any $\mathcal{A} \in \mathbb{F}_{t_0,T}$, $\chi(\mathcal{A}) \triangleq P\{\omega\colon X(\cdot, \omega) \in \mathcal{A}\}$.

Below we introduce conditions (conditions (L1)–(L6)) assuring the existence of realization of the motion produced by the public-signal profile of stochastic strategies in the case of stepwise strategies.

Now let us turn to the definition of an approximate Nash equilibrium. Recall that the Nash equilibrium means that every unilateral changing of strategies does lead to the increasing of the outcome. Thus, we are to introduce the concept of unilateral deviation from the public-signal profile of strategies. It is reasonable to assume that the deviating player has an access to the public signal and can produce his own stochastic signal. This leads us to the following definition.

Definition 3. Given a profile of public-signal correlated strategies

$$\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, P_{x(\cdot)}, u_{x(\cdot)}, v_{x(\cdot)}),$$

we say that a profile of strategies $\mathfrak{w}^c = (\Omega^c, \mathcal{F}^c, \{\mathcal{F}_t^c\}_{t\in[t_0,T]}, P^c_{x(\cdot)}, u^c_{x(\cdot)}, v^c_{x(\cdot)})$ is an unilateral deviation by the first (respectively, the second) player if there exists a filtered measurable space $(\Omega', \mathcal{F}', \{\mathcal{F}'\}_{t\in[t_0,T]})$ such that

(i) $\Omega^c = \Omega \times \Omega'$;

(ii) $\mathcal{F}^c = \mathcal{F} \otimes \mathcal{F}'$;

(iii) $\mathcal{F}_t^c = \mathcal{F}_t \otimes \mathcal{F}'_t$ for $t \in [t_0, T]$;

(iv) for any $x(\cdot) \in C([t_0, T]; \mathbb{R}^d)$ and any $A \in \mathcal{F}$, $P^c_{x(\cdot)}(A \times \Omega') = P_{x(\cdot)}(A)$;

(v) for any $x(\cdot) \in C([t_0, T]; \mathbb{R}^d)$, $t \in [t_0, T]$, $\omega \in \Omega$, $\omega' \in \Omega'$, $v^c_{x(\cdot)}(t, \omega, \omega') = v_{x(\cdot)}(t, \omega)$ (respectively, $u^c_{x(\cdot)}(t, \omega, \omega') = u_{x(\cdot)}(t, \omega)$).

Using this concept of unilateral changing of strategies we receive the following definition of an approximate Nash equilibrium.

For a given initial position (t_0, x_0) and a profile of public-signal correlated strategies $\mathfrak{w}$, we can introduce upper and lower player's outcomes by the following rules:

$$J_i^+(t_0, x_0, \mathfrak{w}) \triangleq \sup\{\mathbb{E}\gamma_i(X(T))\colon (X(\cdot), P) \text{ generated by } \mathfrak{w} \text{ and } (t_0, x_0)\},$$
$$J_i^-(t_0, x_0, \mathfrak{w}) \triangleq \inf\{\mathbb{E}\gamma_i(X(T))\colon (X(\cdot), P) \text{ generated by } \mathfrak{w} \text{ and } (t_0, x_0)\}.$$

Here $\mathbb{E}$ denotes the expectation according to the probability P.

Definition 4. We say that a profile of public-signal correlated strategies $\mathfrak{w}^*$ is a public-signal correlated ε-equilibrium at the position $(t_0, x_0) \in [0, T] \times \mathbb{R}^d$ if, for any profile of strategies $\mathfrak{w}^i$ that is an unilateral deviation from $\mathfrak{w}^*$ by the player i, the following inequality holds true:

$$J_i^+(t_0, x_0, \mathfrak{w}^i) \le J_i^-(t_0, x_0, \mathfrak{w}^*) + \varepsilon.$$

In [2] the approximate Nash equilibria were constructed based on solutions of continuous time stochastic games with dynamics determined by generators of the Lévy–Khintchine type. The general theory of such systems is presented in [13]. In the following, $\mathcal{D}$ stands for a linear subspace of $C^2(\mathbb{R}^d)$ containing $C_b^2(\mathbb{R}^d)$, linear functions $x \mapsto \langle a, x\rangle$ and quadratic functions $x \mapsto \|x - a\|^2$. Let $\Lambda_t[u, v]$ be an operator from $\mathcal{D}$ to $C(\mathbb{R}^d)$ of the form

$$(\Lambda_t[u,v]\phi)(x) \triangleq \frac{1}{2}\langle G(t,x,u,v)\nabla, \nabla\rangle\phi(x) + \langle b(t,x,u,v), \nabla\rangle\phi(x) + \\ + \int_{\mathbb{R}^d} [\phi(x+y) - \phi(x) - \langle y, \nabla\phi(x)\rangle \mathbf{1}_{B_1}(y)]\nu(t,x,u,v,dy). \quad (1.2)$$

Here B_1 stands for the ball of radius 1 centered at the origin; for each $t \in [0, T]$, $x \in \mathbb{R}^d$, $u \in U$, $v \in V$, $G(t, x, u, v)$ is a nonnegative symmetric $d \times d$-matrix, $b(t, x, u, v)$ is a d-dimensional vector, $\nu(t, x, u, v, \cdot)$ is a measure on $\mathbb{R}^d$ such that $\nu(t, x, u, v, \{0\}) = 0$. In the following we call the operator $\Lambda_t[u, v]$ a *generator* [13].

Let us notice that the Markov chain with the Kolmogorov matrix $Q_{x,y}(t,u,v)$, defined for all x, y from at most countable set $\mathcal{S} \subset \mathbb{R}^d$, corresponds to the generator

$$\Lambda_t[u,v]\phi(x) = \sum_{z \in \mathcal{S}} \phi(z) Q_{x,z}(t,u,v),$$

i. e.,

$$\nu(t,x,u,v,A) = \sum_{y \in \mathcal{S} \cap A, \quad y \neq x} Q_{x,x+y}(t,u,v).$$

Here we, without loss of generality, can use scaling and assume that $\|x - y\| \geq 1$, for any $x, y \in \mathcal{S}$, $x \neq y$.

Remark 1. The dynamics (1.1) corresponds to the generator $\Lambda_t[u,v]$ such that, for $\phi \in \mathcal{D}$,

$$\Lambda_t[u,v]\phi(x) = \langle f_1(t,x,u) + f_2(t,x,v), \nabla\phi(x)\rangle.$$

Given a control of players, the generator Λ_t produces the motion that is a stochastic process. In the following we consider the relaxed controls of both players. It is a stochastic process with the values in the set of probabilities on $U \times V$ denoted by $\mathrm{rpm}(U \times V)$. Note that any metric space Υ is naturally embedded into the set of probabilities on Υ by the Dirac measure. Furthermore, the set of probabilities on Υ is compact within the topology of narrow convergence whenever the space Υ is compact.

We use the following definition of control process going back to [8, 10].

Definition 5. Let $s, r \in [0,T]$, $s < r$. We say that a 6-tuple $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t \in [s,r]}, P, \eta, X)$ is a controlled system on $[s,r]$ admissible for the generator $\Lambda_t[u,v]$ if the following conditions hold:

(i) $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [s,r]}, P)$ is a filtered probability space;

(ii) η is a $\{\mathcal{F}_t\}_{t \in [s,r]}$-progressively measurable stochastic process taking values in $\mathrm{rpm}(U \times V)$;

(iii) X is a $\{\mathcal{F}_t\}_{t \in [s,r]}$-adapted stochastic process taking values in $\mathbb{R}^d$;

(iv) for any $\phi \in \mathcal{D}$, the process

$$t \mapsto \phi(X(t)) - \int_s^t \int_{U \times V} (\Lambda_\tau[u,v]\phi)(X(\tau))\eta(\tau, d(u,v))d\tau$$

is a $\{\mathcal{F}_t\}_{t \in [s,r]}$-martingale.

It is assumed that in the auxiliary game with the dynamics determined by Λ the players tries to maximize the values

$$\mathbb{E}\left[\gamma_i(X(T)) + \int_{t_0}^T h_i(t, X(t), u(t), v(t))dt\right].$$

We consider the solution concept for the auxiliary game given by the following condition that is an analog of stability condition first introduced by Krasovskii and Subbotin for the zero-sum differential games [15].

Definition 6. Let $c_1, c_2 : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be continuous functions. We say that the pair (c_1, c_2) satisfies *Condition* $(\mathcal{C})$ if, for any $s, r \in [0,T]$, $s < r$, there exists a filtered measurable space $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}_t^{s,r}\}_{t \in [s,r]})$ satisfying the following properties:

(i) given $y \in \mathbb{R}^d$, one can find processes $\eta_y^{s,r}$, $\widehat{Y}_y^{s,r}$ and a probability $\widehat{P}_y^{s,r}$ such that the 6-tuple $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}_t^{s,r}\}_{t\in[s,r]}, \widehat{P}_y^{s,r}, \eta_y^{s,r}, \widehat{Y}_y^{s,r})$ is a control system admissible for $\Lambda_t[u,v]$ and, for $i = 1, 2$,

$$\widehat{\mathbb{E}}_y^{s,r}\left[c_i(r, \widehat{Y}_y^{s,r}(r)) + \int_s^r \int_{U\times V} h_i(t, \widehat{Y}_y^{s,r}(t), u, v)\eta_y^{s,r}(t, d(u,v))dt\right] = c_i(s,y);$$

(ii) for any $y \in \mathbb{R}^d$ and $v \in V$, one can find a relaxed stochastic control of the first player $\mu_{y,v}^{s,r}$, a process $\overline{Y}_{y,v}^{1,s,r}$ taking values in $\mathbb{R}^d$ and a probability $\overline{P}_{y,v}^{1,s,r}$ such that the 6-tuple $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}_t^{s,r}\}_{t\in[s,r]}, \overline{P}_{y,v}^{1,s,r}, \mu_{y,v}^{s,r}\otimes\delta_v, \overline{Y}_{y,v}^{1,s,r})$ is a control system admissible for $\Lambda_t[u,v]$ and

$$\overline{\mathbb{E}}_{y,v}^{1,s,r}\left[c_2(r, \overline{Y}_{y,v}^{1,s,r}(r)) + \int_s^r \int_U h_2(t, \overline{Y}_{y,v}^{1,s,r}(t), u, v)\mu_{y,v}^{s,r}(t, du)dt\right] \le c_2(s,y);$$

(iii) given $y \in \mathbb{R}^d$ and $u \in U$, one can find a second player's relaxed stochastic control $\nu_{y,u}^{s,r}$, a process $\overline{Y}_{y,u}^{2,s,r}$ and a probability $\overline{P}_{y,u}^{2,s,r}$ such that the 6-tuple

$$(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}_t^{s,r}\}_{t\in[s,r]}, \overline{P}_{y,u}^{2,s,r}, \delta_u\otimes\nu_{y,u}^{s,r}, \overline{Y}_{y,u}^{2,s,r})$$

is a control system admissible for $\Lambda_t[u,v]$ and

$$\overline{\mathbb{E}}_{y,u}^{2,s,r}\left[c_1(r, \overline{Y}_{y,u}^{2,s,r}(r)) + \int_s^r \int_V h_1(t, \overline{Y}_{y,u}^{2,s,r}(t), u, v)\nu_{y,u}^{s,r}(t, dv)dt\right] \le c_1(s,y).$$

Here $\widehat{\mathbb{E}}_y^{s,r}$ (respectively, $\overline{\mathbb{E}}_{y,u}^{1,s,r}$, $\overline{\mathbb{E}}_{y,u}^{2,s,r}$) denotes the expectation according to the probability $\widehat{P}_y^{s,r}$ (respectively, $\overline{P}_{y,u}^{1,s,r}$, $\overline{P}_{y,u}^{2,s,r}$).

Let us comment on this condition. The pair of functions (c_1, c_2) is an analog of the value function for the auxiliary games. The first part means that both players can maintain the values of the functions c_1, c_2 through some trajectory. The second (respectively, third) condition states that the first (respectively, second) player can punish his partner if he plays with the constant strategy. These conditions is a stochastic version of the condition introduced in [1]. That condition provides the existence of a universal Nash equilibrium in the class of strategies with guide.

To construct an approximate Nash equilibrium let us denote

$$\Sigma(t,x,u,v) \triangleq \mathrm{tr}G(t,x,u,v) + \int_{\mathbb{R}^d} \|y\|^2\nu(t,x,u,v,dy), \tag{1.3}$$

$$g(t,x,u,v) \triangleq b(t,x,u,v) + \int_{\mathbb{R}^d\setminus B_1} y\nu(t,x,u,v,du). \tag{1.4}$$

The value Σ estimates the randomness of the dynamics determined by the generator Λ, whereas g is an effective drift in the auxiliary game.

The following assumptions are imposed.

(L1) U, V are metric compacts;

(L2) the functions f_1, f_2 G, b, γ_1, γ_2, h_1, h_2 are continuous and bounded;

(L3) for any $\phi \in \mathcal{D}$, the function $[0,T]\times\mathbb{R}^d\times U\times V \ni (t,x,u,v) \mapsto \int_{\mathbb{R}^d}\phi(y)\nu(t,x,u,v,dy)$ is continuous.

(L4) there exists a function $\alpha(\cdot) : \mathbb{R} \to [0, +\infty)$ such that $\alpha(\delta) \to 0$ as $\delta \to 0$ and, for any $t, s \in [0, T]$, $x \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$\|f(t, x, u, v) - f(s, x, u, v)\| \leq \alpha(t - s),$$

$$\|g(t, x, u, v) - g(s, x, u, v)\| \leq \alpha(t - s);$$

(L5) there exists a constant M such that, for any $t \in [0, T]$, $x \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$\|f(t, x, u, v)\| \leq M, \quad \|g(t, x, u, v)\| \leq M;$$

(L6) there exists a constant $K > 0$ such that, for any $t \in [0, T]$, $x', x'' \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$\|f(t, x', u, v) - f(t, x'', u, v)\| \leq K\|x' - x''\|,$$

$$\|g(t, x', u, v) - g(t, x'', u, v)\| \leq K\|x' - x''\|;$$

(L7) there exists a constant $R > 0$ such that, for any $x', x'' \in \mathbb{R}^d$, $i = 1, 2$,

$$|\gamma_i(x') - \gamma_i(x'')| \leq R\|x' - x''\|;$$

(L8) for any $t \in [0, T]$, $x \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$|\Sigma(t, x, u, v)| \leq \delta^2,$$

$$\|f(t, x, u, v) - g(t, x, u, v)\|^2 \leq 2\delta^2,$$

$$|h_i(t, x, u, v)| \leq \delta.$$

In condition (L8) δ is a small parameter.

Let us briefly comment on the imposed conditions. First, recall that the function g plays the role of an effective drift, whereas Σ is an analog of the squared violence coefficient. Condition (L1) is rather standard in the theory of differential games. Other conditions provide the continuity properties. We assume that the dynamics of the original function and the effective drift are both uniformly continuous w.r.t. time, Lipschitz continuous w.r.t. phase variable and bounded. Additionally, we assume Lipschitz continuity of the payoff function (condition (L7)). Finally, condition (L8) states that the original and auxiliary games are close.

Now let us discuss the existence of the motion. The public-signal profile of stochastic strategies $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, P_{x(\cdot)}, u_{x(\cdot)}, v_{x(\cdot)})$ is called stepwise, if there exists a partition $\Delta = \{t_i\}_{i=0}^N$ of the time interval $[t_0, T]$ such that, for any $n = 0, \ldots, N - 1$, and every $x(\cdot), y(\cdot) \in C([t_0, T]; \mathbb{R}^d)$ the condition $x(t_i) = y(t_i)$, $i = 0, \ldots, n$ implies that $P_{x(\cdot)}(A) = P_{y(\cdot)}(A)$, when $A \in \mathcal{F}_t$ and $u_{x(\cdot)}(t) = u_{y(\cdot)}(t)$ $P_{x(\cdot)} - a.s.$ when $t \in [t_0, t_{n+1})$. It can be proved that under conditions (L1)–(L8), given a stepwise public-signal profile of stochastic strategies $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, P_{x(\cdot)}, u_{x(\cdot)}, v_{x(\cdot)})$, there exists at least one realization of the motion generated by $\mathfrak{w}$ and initial position (t_0, x_0). The proof is by induction.

Now let us present the theorem proved in [2] providing the existence of an approximate Nash equilibrium and the estimates of the approximation rate. Set

$$\beta \triangleq (5 + 2K), \tag{1.5}$$

$$C \triangleq 2\sqrt{Te^{\beta T}}. \tag{1.6}$$

Theorem 1. Let continuous functions $c_1, c_2 : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ be such that

- $c_i(T, x) = \gamma_i(x)$;
- (c_1, c_2) satisfies Condition $(\mathcal{C})$.

Then, for any $(t_0, x_0) \in [0, T] \times \mathbb{R}^d$, and $\varepsilon > (RC + T)\delta$, there exists a profile of public-signal correlated strategies $\mathfrak{w}^*$ that is ε-equilibrium at (t_0, x_0). Moreover, if X^* and P^* are generated by $\mathfrak{w}^*$ and (t_0, x_0), $\mathbb{E}^*$ denotes the expectation according to P^*, then

$$|\mathbb{E}^*\gamma_i(X^*(T)) - c_i(t_0, x_0)| \leq \varepsilon.$$

§ 2 Markov games

In this section we examine a nonzero-sum two-player continuous time Markov game. We will show that the solution of the Bellman system for the Markov game satisfies condition $(\mathcal{C})$.

Let $\mathcal{S}$ be at most countable set of states. We assume that there exist Kolmogorov matrices $Q^1_{x,y}(t, u)$, $Q^2_{x,y}(t, v)$ satisfying the following property: for some natural number l, given $x \in \mathcal{S}$, there exist $Y(x) \subset \mathcal{S}$ such that $|Y(x)| = l$ and $Q^1_{x,y}(t, u) = Q^2_{x,y}(t, v) = 0$ when $y \notin Y(x)$. Notice that this Markov chain corresponds to the generator

$$\Lambda_t[u, v]\phi(x) = \sum_{z\in\mathcal{S}} \phi(z)[Q^1_{x,z}(t, u) + Q^2_{x,z}(t, v)]. \tag{2.1}$$

We assume that the first player tries to maximize the value

$$\mathbb{E}\left[\gamma_1(x(T)) + \int_0^T h_1(t, x(t), u(t))dt\right],$$

when the second player wishes to maximize

$$\mathbb{E}\left[\gamma_2(x(T)) + \int_0^T h_2(t, x(t), v(t))dt\right].$$

Further, let us introduce an analog of Bellman equation for the controlled Markov chain. To this end, given $t \in [0, T]$, $x \in \mathcal{S}$, $\xi : \mathcal{S} \to \mathbb{R}$, $\mu \in \mathcal{P}(U)$, $\nu \in \mathcal{P}(V)$ set

$$\widehat{H}_1(t, x, \xi, \mu) \triangleq \int_U \left[\sum_{y\in\mathcal{S}} Q^1_{x,y}(t, x, u)\xi(y) + h_1(t, x, u)\right] \mu(du),$$

$$\widehat{H}_2(t, x, \xi, \nu) \triangleq \int_V \left[\sum_{y\in\mathcal{S}} Q^2_{x,y}(t, x, v)\xi(y) + h_2(t, x, u)\right] \nu(du).$$

The functions $\widehat{H}_1$, $\widehat{H}_2$ play the role of pre-Hamiltonians. Further, put

$$\mathcal{O}_1(t, x, \xi) \triangleq \underset{\mu\in\mathcal{P}(U)}{\mathrm{Argmax}} \widehat{H}_1(t, x, \xi, \mu), \quad \mathcal{O}_2(t, x, \xi) \triangleq \underset{\nu\in\mathcal{P}(V)}{\mathrm{Argmax}} \widehat{H}_2(t, x, \xi, \nu).$$

Finally, if ξ_1, ξ_2 are real valued functions defined on $\mathcal{S}$, denote

$$\mathcal{H}_1(t, x, \xi_1, \xi_2) \triangleq \max_{\mu\in\mathcal{P}(U)} \widehat{H}_1(t, x, \xi_1, \mu) + \left\{\int_V \sum_{y\in\mathcal{S}} Q^2_{x,y}(t, x, v)\xi(y)\nu(du) : \nu \in \mathcal{O}_2(t, x, \xi_2)\right\},$$

$$\mathcal{H}_2(t, x, \xi_1, \xi_2) \triangleq \max_{\nu\in\mathcal{P}(V)} \widehat{H}_2(t, x, \xi_2, \nu) + \left\{\int_U \sum_{y\in\mathcal{S}} Q^1_{x,y}(t, x, u)\xi(y)\mu(du) : \mu \in \mathcal{O}_1(t, x, \xi_1)\right\}.$$

The multifunctions $\mathcal{H}_1$ and $\mathcal{H}_2$ are analogs of Hamiltonians.

Below we consider the following differential inclusions which are natural analogs of Bellman equation

$$\frac{d}{dt}\xi_i(t,x) \in -\mathcal{H}_i(t,x,\xi_1,\xi_2), \quad \xi_i(T,x) = \gamma_i(x), \quad i = 1,2. \tag{2.2}$$

Theorem 2. Assume that $Q^1(t,u)$ and $Q^2(t,v)$ are continuous. Then, if the pair $(\xi_1(\cdot,\cdot),\xi_2(\cdot,\cdot))$ solves (2.2), then it satisfies condition $(\mathcal{C})$.

Proof. Let $s,r \in [0,T]$, $s<r$. First, set

$$\widehat{\Omega}^{s,r} \triangleq D([s,r];\mathcal{S}).$$

Here $D([s,r],\mathcal{S})$ stands for the set of càdlàg functions with values in $\mathcal{S}$. To define a filtration, for $t_1,\ldots,t_k \in [s,r]$, $C_1,\ldots,C_k \subset \mathcal{S}$, denote by $A^{s,r}_{t_1,\ldots,t_k,C_1,\ldots,C_k}$ the set of càdlàg functions $x(\cdot) \in D([s,r];\mathcal{S})$ such that, for each t_i, $x(t_i) \in C_i$. Now, put

$$\widehat{\mathcal{F}}^{s,r}_t \triangleq \mathcal{B}(\{A^{s,r}_{t_1,\ldots,t_k;C_1,\ldots,C_k} : t_1,\ldots,t_k \in [s,t]\}),$$

$$\widehat{\mathcal{F}}^{s,r} \triangleq \widehat{\mathcal{F}}^{s,r}_r.$$

Further, let $\hat{\mu}$ be a weakly measurable function from $[0,T]\times\mathcal{S}$ to $\mathcal{P}(U)$ such that

$$\hat{\mu}(t,x) \in \mathcal{O}_1(t,x,\xi_1(\cdot,\cdot)).$$

Similarly, pick a weakly measurable function $\hat{\nu} : [0,T]\times\mathcal{S} \to \mathcal{P}(V)$ satisfying

$$\hat{\nu}(t,x) \in \mathcal{O}_2(t,x,\xi_2(\cdot,\cdot)).$$

If y_0 is a given initial position, then let $\widehat{P}^{s,r}_{y_0}$ be a distribution of paths for the Markov chain with the Kolmogorov matrix

$$\widehat{Q}_{x,y}(t) \triangleq \int_U Q^1_{x,y}(t,u)\hat{\mu}(t,x,du) + \int_V Q^2_{x,y}(t,v)\hat{\nu}(t,x,dv)$$

and the initial distribution equal to δ_{y_0}. Here δ_{y_0} stands for the Dirac measure concentrated at y_0. Now, define the process $\widehat{Y}^{s,r}_{y_0}$ by the rule, for $t \in [s,r]$, $\omega \in \widehat{\Omega}^{s,r} = D([s,r],\mathcal{S})$,

$$\widehat{Y}^{s,r}_{y_0}(t,\omega) \triangleq \omega(t).$$

Finally, put

$$\eta^{s,r}_{y_0}(t,\omega,d(u,v)) \triangleq \hat{\mu}(t,\omega(t),du) \otimes \hat{\nu}(t,\omega(t),dv).$$

By construction $(\widehat{\Omega}^{s,r},\widehat{\mathcal{F}}^{s,r},\widehat{\mathcal{F}}^{s,r}_t,\widehat{P}^{s,r}_{y_0},\eta^{s,r},\widehat{Y}^{s,r}_{y_0})$ is a control process admissible for the generator $\Lambda_t[u,v]$ given by (1.2).

This implies the equality

$$\widehat{E}^{s,r}_{y_0}\xi_1(r,\widehat{Y}^{s,r}_{y_0}(r)) - \xi_1(s,y_0) = \widehat{E}^{s,r}_{y_0}\int_s^r \dot{\xi}_1(t,\widehat{Y}^{s,r}_{y_0}(t))dt +$$
$$+\widehat{E}^{s,r}_{y_0}\int_s^r\int_U\int_V\sum_{z\in\mathcal{S}}\xi_1(t,z)\left[Q^1_{\widehat{Y}^{s,r}_{y_0}(t),z}(t,u) + Q^2_{\widehat{Y}^{s,r}_{y_0}(t),z}(t,v)\right]\hat{\mu}(t,\widehat{Y}^{s,r}_{y_0}(t),du)\hat{\nu}(t,\widehat{Y}^{s,r}_{y_0}(t),dv)\,dt.$$

Using the definition of $\mathcal{H}_1$, and the fact that (ξ_1, ξ_2) solves (2.2), we get

$$\widehat{E}^{s,r}_{y_0}\xi_1(r, \widehat{Y}^{s,r}_{y_0}(r)) - \xi_1(s, y_0) = -\widehat{E}^{s,r}_{y_0}\int_s^r\int_U [h_1(t, \widehat{Y}^{s,r}_{y_0}(t), u)]\hat{\mu}(t, \widehat{Y}^{s,r}_{y_0}(t), du)dt.$$

Analogously, one can prove that

$$\widehat{E}^{s,r}_{y_0}\xi_2(r, \widehat{Y}^{s,r}_{y_0}(r)) - \xi_2(s, y_0) = -\widehat{E}^{s,r}_{y_0}\int_s^r\int_V [h_2(t, \widehat{Y}^{s,r}_{y_0}(t), v)]\hat{\nu}(t, \widehat{Y}^{s,r}_{y_0}(t), dv)dt.$$

This proves part (i) of Condition $(\mathcal{C})$.

To prove the second part, pick $v \in V$ and consider the probability on distribution of paths $\overline{P}^{1,s,r}_{y_0,v}$ produced by the Markov chain with the Kolmogorov matrix

$$\overline{Q}^{1,v}_{x,y}(t) \triangleq \int_U Q^1_{x,y}(t, u)\hat{\mu}(t, x, du) + Q^2_{x,y}(t, v)$$

and initial distribution equal to δ_{y_0}. Put

$$\overline{Y}^{1,s,r}_{y,v}(t, \omega) \triangleq \omega(t), \quad \mu^{s,r}_{y,v}(t, \omega, du) \triangleq \hat{\mu}(t, \omega(t), du).$$

Notice that $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \widehat{\mathcal{F}}^{s,r}_t, \overline{P}^{1,s,r}_{y_0,v}, \mu^{s,r}_{y,v}\otimes\delta_v, \overline{Y}^{1,s,r}_{y_0,v})$ is a control process admissible for the generator $\Lambda_t[u, v]$ given by (1.2). As above, we have that

$$\begin{aligned}
&\overline{E}^{1,s,r}_{y_0,v}\xi_2(r, \overline{Y}^{1,s,r}_{y_0,v}(r)) - \xi_2(s, y_0)\\
&\quad= \overline{E}^{1,s,r}_{y_0,v}\int_s^r \dot{\xi}_2(t, \overline{Y}^{1,s,r}_{y_0,v}(t))dt\\
&\qquad+ \overline{E}^{1,s,r}_{y_0,v}\int_s^r\int_U\sum_{z\in\mathcal{S}}\xi_2(t, z)\left[Q^1_{\overline{Y}^{1,s,r}_{y_0,v}(t),z}(t, u) + Q^2_{\overline{Y}^{1,s,r}_{y_0,v}(t),z}(t, v)\right]\hat{\mu}(t, \overline{Y}^{1,s,r}_{y_0,v}(t), du)dt
\end{aligned}$$

Since

$$\int_U\Big[\sum_{z\in\mathcal{S}}\xi_2(t, z)\left[Q^1_{x,z}(t, u) + Q^2_{x,z}(t, v)\right] + h_2(t, x, v)\Big]\hat{\mu}(t, x, du) \le \mathcal{H}_2(t, x, \xi_1, \xi_2),$$

using the assumption that the pair (ξ_1, ξ_2) is a solution of (2.2), we conclude that

$$\overline{E}^{1,s,r}_{y_0,v}\xi_2(r, \overline{Y}^{1,s,r}_{y_0,v}(r)) - \xi_2(s, y_0) \le -\overline{E}^{1,s,r}_{y_0,v}\int_s^r [h_2(t, \overline{Y}^{1,s,r}_{y_0,v}(t), v)]dt.$$

This shows part (ii) of Condition $(\mathcal{C})$. The third part is proved in the same way. □

§ 3 Construction of approximating Markov game

The aim of this section is to present an example of Markov games approximating the original nonzero-sum differential game with the dynamics given by (1.1), where player i tries to maximize $\gamma_i(x(T))$.

Let us rewrite dynamics (1.1) in a coordinate-wise form:

$$\frac{d}{dt}x_j(t) = f_{1,j}(t, x_1(t), \ldots, x_d(t), u) + f_{2,j}(t, x_1(t), \ldots, x_d(t), v).$$

Here $x_j(t)$ stands for the j-th coordinate of the vector $x(t)$, whilst

$$f_1(t, x, u) = (f_{1,1}(t, x, u), \ldots, f_{1,d}(t, x, u)),$$

$$f_2(t,x,v) = (f_{2,1}(t,x,u),\dots,f_{2,d}(t,x,v)).$$

Let $\varkappa$ be a positive number. Further, denote by e^j the j-th coordinate vector.

Put

$$Q^1_{x,y}(t,u) \triangleq \begin{cases} \frac{1}{\varkappa}|f_{1,j}(t,x,u)|, & y = x + \varkappa\,\mathrm{sgn}(f_{1,j}(t,x,u))\cdot e^j,\\ -\frac{1}{\varkappa}\sum_{j=1}^d |f_{1,j}(t,x,u)|, & y = x\\ 0, & \text{otherwise.}\end{cases}$$

Analogously, set

$$Q^2_{x,y}(t,v) \triangleq \begin{cases} \frac{1}{\varkappa}|f_{2,j}(t,x,u)|, & y = x + \varkappa\,\mathrm{sgn}(f_{2,j}(t,x,v))\cdot e^j,\\ -\frac{1}{\varkappa}\sum_{j=1}^d |f_{2,j}(t,x,v)|, & y = x\\ 0, & \text{otherwise.}\end{cases}$$

Notice that if we consider the Markov chain with the Kolmogorov matrix $Q_{x,y}(t,u,v) \triangleq Q^1_{x,y}(t,u) + Q^2_{x,y}(t,v)$ and assume that the initial state lies at $h\mathbb{Z}^d$, then the state of the Markov chain will always belong to $\varkappa\mathbb{Z}^d$. Thus, we can put $\mathcal{S} \triangleq \varkappa\mathbb{Z}^d$.

Now let us compute g and Σ. We have that

$$\nu(t,x,u,v,du) = \sum_{j=1}^d |f_{1,j}(t,x,u)|\delta_{(\varkappa\,\mathrm{sgn}(f_{1,j}(t,x,u)))e^j}(dy) \\ + \sum_{j=1}^d |f_{2,j}(t,x,v)|\delta_{(\varkappa\,\mathrm{sgn}(f_{2,j}(t,x,v)))e^j}(dy).$$

Thus, we have that

$$g(t,x,u,v) = f(t,x,u,v), \tag{3.1}$$

whereas

$$\Sigma(t,x,u,v) = \varkappa\sum_{j=1}^d(|f_1(t,x,u)| + |f_2(t,x,v)|) \leq 2M\varkappa. \tag{3.2}$$

This implies that if $f_1(t,x,u)+f_2(t,x,v)$ is Lipschitz continuous and bounded by M, then the proposed Markov approximation provides the stochastic system satisfying conditions (L1)–(L8).

Notice that

$$(\xi Q^1(t,u))(x) = \sum_{j=1}^d \frac{1}{\varkappa}|f_{1,j}(t,x,u)|(\xi(x+\varkappa\,\mathrm{sgn}(f_{1,j}(t,x,u))) - \xi(x)),$$

$$(\xi Q^2(t,v))(x) = \sum_{j=1}^d \frac{1}{\varkappa}|f_{2,j}(t,x,v)|(\xi(x+\varkappa\,\mathrm{sgn}(f_{2,j}(t,x,v))) - \xi(x)).$$

This provides the precise formulas for H_i and for $\mathcal{H}_i$ for any appropriate choice of functions h_1, h_2. One can plug them into inclusion 2.2. Any solution of this inclusion (ξ_1,ξ_2) satisfies condition $(\mathcal{C})$ (see Theorem 2) and, thus, provides an approximate Nash equilibrium by Theorem 1. If $\|h_1\|,\|h_2\| \leq \sqrt{\varkappa}$, the approximation rate is of order $\sqrt{\varkappa}$ (see condition (L8), (3.1), (3.2)).

Notice that we have considered the case when $\mathcal{S}$ is equal to $\varkappa\mathbb{Z}^d$ and, thus, countable. However, one can easily construct the Markov approximation with the finite state space. To this end assume that we are interested in the value function only in the subset of state space $A \subset \mathbb{R}^d$.

Since f is bounded by M, we have that every trajectory started at A does not leave the set $A_1 \triangleq A + B_{MT}$, where B_r stands for the ball of radius r centered at the origin. Further, let $A_2 \triangleq A_1 + B_{M/K}$. For $x \in A_1 \cup (\mathbb{R}^d \setminus A_2)$, we define the functions

$$f_1'(t,x,u) \triangleq \begin{cases} f_1(t,x,u), & x \in A_1, \\ 0 & x \notin \mathbb{R}^d \setminus A_2, \end{cases} \qquad f_2'(t,x,v) \triangleq \begin{cases} f_2(t,x,v), & x \in A_1, \\ 0 & x \notin \mathbb{R}^d \setminus A_2, \end{cases}$$

The function $f'(t,x,u,v) \triangleq f_1'(t,x,u) + f_2'(t,x,v)$ is continuous, bounded by M, uniformly continuous w. r. t. time, Lipschitz continuous w. r. t. x with the constant K and equal to zero on $\mathbb{R}^d \setminus A_2$. It can be extended to the whole space with the same Lipshitz constant and the same modulus of continuity w. r. t. to t. Since $f'(t,x,u,v) = f(t,x,u,v)$ for $x \in A_1$, we have that the solutions of the nonzero-sum game for the original dynamics and the dynamics given by

$$\frac{d}{dt}x(t) = f'(t,x(t),u,v)$$

coincide for $x \in A$. On the other hand, we can consider the Markov game for the function f' only for the set $B_2^{\varkappa} \triangleq A + B_{MT+M/K+M\varkappa} \cap \varkappa\mathbb{Z}^d$ due to the fact that outside this state the dynamics is equal to zero.

§ 4 Conclusion

The paper deals with the Markov approximations of the nonzero-sum differential game. The cornerstone of our consideration is the general result proved in [2] stating that, given an auxiliary continuous-time stochastic game, and a pair of functions satisfying stability condition $(\mathcal{C})$ for this auxiliary game, one can construct an approximate Nash equilibrium for the original game. The principal example here is Markov game, i. e., the game with the dynamics determined by a continuous-time Markov chain. We write down the system of differential inclusions playing the role of the system of Bellman equations for the Markov game and prove that every its solution satisfies condition $(\mathcal{C})$. Finally, we introduce the method of construction of a Markov game by the given differential game and estimate the approximation rate for the corresponding Nash equilibrium.

Funding. This work was funded by the Russian Science Foundation (Project No. 17-11-01093).

REFERENCES

1. Averboukh Y. Universal Nash equilibrium strategies for differential games, *Journal of Dynamical and Control Systems*, 2015, vol. 21, issue 3, pp. 329–350. https://doi.org/10.1007/s10883-014-9224-9
2. Averboukh Y. Approximate public-signal correlated equilibria for nonzero-sum differential games, *SIAM Journal on Control and Optimization*, 2019, vol. 57, issue 1, pp. 743–772. https://doi.org/10.1137/17M1161403
3. Başar T., Olsder G. *Dynamic noncooperative game theory*, Philadelphia: SIAM, 1998.
4. Bressan A., Shen W. Semi-cooperative strategies for differential games, *International Journal of Game Theory*, 2004, vol. 32, issue 4, pp. 561–593. https://doi.org/10.1007/s001820400180
5. Bressan A., Shen W. Small BV solutions of hyperbolic noncooperative differential games, *SIAM Journal on Control and Optimization*, 2004, vol. 43, issue 1, pp. 194–215. https://doi.org/10.1137/S0363012903425581
6. Cardaliaguet P., Plaskacz S. Existence and uniqueness of a Nash equilibrium feedback for a simple nonzero-sum differential game, *International Journal of Game Theory*, 2003, vol. 32, issue 1, pp. 33–71. https://doi.org/10.1007/s001820300152

7. Chistyakov S. V. On coalition-free differential games, *Soviet Mathematics. Doklady*, 1981, vol. 24, pp. 166–169.
8. Fleming W. H., Soner H. M. *Controlled Markov processes and viscosity solutions*, New York: Springer, 2006. https://doi.org/10.1007/0-387-31071-1
9. Friedman A. *Differential games*, New York: Dover Publications, 2013.
10. Gihman I. I., Skorohod A. V. *Controlled stochastic processes*, New York: Springer, 1979. https://doi.org/10.1007/978-1-4612-6202-2
11. Hamadène S., Mannucci P. Regularity of Nash payoffs of Markovian nonzero-sum stochastic differential games, *Stochastics*, 2019, vol. 91, issue 5, pp. 695–715. https://doi.org/10.1080/17442508.2018.1540627
12. Kleimenov A. *Neantagonisticheskie pozitsionnye differentsial'nye igry* (Non-antagonistic positional differential games), Yekaterinburg: Nauka, 1993.
13. Kolokoltsov V. N. *Markov processes, semigroups and generators*, Berlin: De Gruyter, 2011.
14. Kononenko A. F. On equilibrium positional strategies in nonantagonistic differential games, *Soviet Mathematics. Doklady*, 1976, vol. 17, pp. 1557–1560.
15. Krasovskii N. N., Subbotin A. I. *Game-theoretical control problems*, New York: Springer, 1988.
16. Levy Y. Continuous-time stochastic games of fixed duration, *Dynamic Games and Applications*, 2013, vol. 3, issue 2, pp. 279–312. https://doi.org/10.1007/s13235-012-0067-2
17. Mannucci P. Nonzero-sum stochastic differential games with discontinuous feedback, *SIAM Journal on Control and Optimization*, vol. 43, issue 4, pp. 1222–1233. https://doi.org/10.1137/S0363012903423715
18. Mannucci P. Nash points for nonzero-sum stochastic differential games with separate Hamiltonians, *Dynamic Games and Applications*, 2014, vol. 4, issue 3, pp. 329–344. https://doi.org/10.1007/s13235-013-0101-z
19. Meyer P. A. *Probability and potentials*, Waltham, Mass.: Blaisdell Publishing Company, 1966.
20. Tolwinski B., Haurie A., Leitman G. Cooperative equilibria in differential games, *Journal of Mathematical Analysis and Applications*, 1986, vol. 119, issues 1–2, pp. 182–202. https://doi.org/10.1016/0022-247X(86)90152-6

Received 17.11.2019

Averboukh Yurii, PhD, Researcher, Department of Control Systems, Krasovskii Institute of Mathematics and Mechanics, 16, ul. S. Kovalevskoi, Yekaterinburg, 620219, Russia.
Institute of Natural Sciences and Mathematics, Ural Federal University, ul. Turgeneva, 4, Yekaterinburg, 620000, Russia.
E-mail: ayv@imm.uran.ru

Citation: Yu. V. Averboukh. Markov approximations of nonzero-sum differential games, *Vestnik Udmurtskogo Universiteta. Matematika. Mekhanika. Komp'yuternye Nauki*, 2020, vol. 30, issue 1, pp. 3–17.

ВЕСТНИК УДМУРТСКОГО УНИВЕРСИТЕТА. МАТЕМАТИКА. МЕХАНИКА. КОМПЬЮТЕРНЫЕ НАУКИ

МАТЕМАТИКА 2020. Т. 30. Вып. 1. С. 3–17.

Ю. В. Авербух

Марковские аппроксимации неантагонистических дифференциальных игр

Ключевые слова: неантагонистические дифференциальные игры, приближенное равновесия по Нэшу, марковские игры, дифференциальные включения.

УДК 517.977.8

DOI: 10.35634/vm200101

В статье рассматриваются приближенные решения неантагонистических дифференциальных игр. Приближенное равновесие по Нэшу может быть построено по заданному решению вспомогательной стохастической игры с непрерывным временем. Мы рассматриваем случай, когда динамика вспомогательной игры задается марковской цепью с непрерывным временем. Для этой игры функция цены определяется решением системы обыкновенных дифференциальных включений. Таким образом, мы получаем конструкцию приближенного равновесия по Нэшу с выигрышами игроков, близкими к решениям системы обыкновенных дифференциальных включений. Также предложен способ построения марковской игры с непрерывным временем, аппроксимирующей исходную игру.

Финансирование. Работа выполнена при поддержке РНФ, проект № 17–11–01093.

СПИСОК ЛИТЕРАТУРЫ

1. Averboukh Y. Universal Nash equilibrium strategies for differential games // Journal of Dynamical and Control Systems. 2015. Vol. 21. Issue 3. P. 329–350. https://doi.org/10.1007/s10883-014-9224-9
2. Averboukh Y. Approximate public-signal correlated equilibria for nonzero-sum differential games // SIAM Journal on Control and Optimization. 2019. Vol. 57. Issue 1. P. 743–772. https://doi.org/10.1137/17M1161403
3. Başar T., Olsder G. Dynamic noncooperative game theory. Philadelphia: SIAM, 1998.
4. Bressan A., Shen W. Semi-cooperative strategies for differential games // International Journal of Game Theory. 2004. Vol. 32. Issue 4. P. 561–593. https://doi.org/10.1007/s001820400180
5. Bressan A., Shen W. Small BV solutions of hyperbolic noncooperative differential games // SIAM Journal on Control and Optimization. 2004. Vol. 43. Issue 1. P. 194–215. https://doi.org/10.1137/S0363012903425581
6. Cardaliaguet P., Plaskacz S. Existence and uniqueness of a Nash equilibrium feedback for a simple nonzero-sum differential game // International Journal of Game Theory. 2003. Vol. 32. Issue 1. P. 33–71. https://doi.org/10.1007/s001820300152
7. Чистяков С. В. О бескоалиционных дифференциальных играх // Доклады АН СССР. 1981. Т. 259. № 5. С. 1052–1055. http://mi.mathnet.ru/dan44642
8. Fleming W. H., Soner H. M. Controlled Markov processes and viscosity solutions. New York: Springer, 2006. https://doi.org/10.1007/0-387-31071-1
9. Friedman A. Differential games. New York: Dover Publications, 2013.
10. Gihman I. I., Skorohod A. V. Controlled stochastic processes. New York: Springer, 1979. https://doi.org/10.1007/978-1-4612-6202-2
11. Hamadène S., Mannucci P. Regularity of Nash payoffs of Markovian nonzero-sum stochastic differential games // Stochastics. 2019. Vol. 91. Issue 5. P. 695–715. https://doi.org/10.1080/17442508.2018.1540627
12. Клейменов А. Ф. Неантагонистические позиционные дифференциальные игры. Екатеринбург: Наука, 1993.
13. Kolokoltsov V. N. Markov processes, semigroups and generators. Berlin: De Gruyter, 2011.
14. Кононенко А. Ф. О равновесных позиционных стратегиях в неантагонистических дифференциальных играх // Доклады АН СССР. 1976. Т. 231. № 2. С. 285–288. http://mi.mathnet.ru/dan40738

15. Krasovskii N. N., Subbotin A. I. Game-theoretical control problems. New York: Springer, 1988.
16. Levy Y. Continuous-time stochastic games of fixed duration // Dynamic Games and Applications. 2013. Vol. 3. Issue 2. P. 279–312. https://doi.org/10.1007/s13235-012-0067-2
17. Mannucci P. Nonzero-sum stochastic differential games with discontinuous feedback // SIAM Journal on Control and Optimization. Vol. 43. Issue 4. P. 1222–1233. https://doi.org/10.1137/S0363012903423715
18. Mannucci P. Nash points for nonzero-sum stochastic differential games with separate Hamiltonians // Dynamic Games and Applications. 2014. Vol. 4. Issue 3. P. 329–344. https://doi.org/10.1007/s13235-013-0101-z
19. Meyer P. A. Probability and potentials. Waltham, Mass.: Blaisdell Publishing Company, 1966.
20. Tolwinski B., Haurie A., Leitman G. Cooperative equilibria in differential games // Journal of Mathematical Analysis and Applications. 1986. Vol. 119. Issues 1–2. P. 182–202. https://doi.org/10.1016/0022-247X(86)90152-6

Поступила в редакцию 17.11.2019

Авербух Юрий Владимирович, к. ф.-м. н., снс, отдел управляемых систем, Институт математики и механики им. Н. Н. Красовского Уральского отделения Российской академии наук, 620219, Россия, г. Екатеринбург, ул. С. Ковалевской, 16.
Институт естественных наук и математики, Уральский федеральный университет, 620000, Россия, г. Екатеринбург, ул. Тургенева, 4.
E-mail: ayv@imm.uran.ru

Цитирование: Ю. В. Авербух. Марковские аппроксимации неантагонистических дифференциальных игр // Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. 2020. Т. 30. Вып. 1. С. 3–17.